

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 9, 2003, 16:24:23 ; Search time 18.4571 Seconds  
(without alignments)  
98.997 Million cell updates/sec

Title: US-09-905-691-3  
Perfect score: 19  
Sequence: 1 AEARARRAARARRAARA 19

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR\_76:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	47.4	850	2 JC5700	ErbB kinase activa
2	9	47.4	860	2 JC5702	ErbB kinase activa
3	9	47.4	868	2 JC5701	ErbB kinase activa
4	8	42.1	144	2 H72702	hypothetical prote
5	8	42.1	284	2 T34685	ABC-transporter At
6	8	42.1	328	2 H83363	pyoverdine biosynt
7	8	42.1	334	2 E95279	probable (imported
8	8	42.1	395	1 K1SMG	galactokinase (EC
9	8	42.1	501	2 T34741	hypothetical prote
10	8	42.1	532	2 E87343	ABC transporter, A
11	8	42.1	2252	2 S06188	genome polyprotein
12	7	36.8	57	2 G84072	hypothetical prote
13	7	36.8	71	2 A13052	hypothetical prote
14	7	36.8	91	2 A87663	hypothetical prote
15	7	36.8	126	2 T52525	tyrosinase co-fact
16	7	36.8	134	2 A26986	hypothetical prote
17	7	36.8	145	2 E75622	hypothetical prote
18	7	36.8	146	2 A23971	hypothetical prote
19	7	36.8	149	2 T39071	60S ribosomal prot
20	7	36.8	149	2 T41324	60S ribosomal prot
21	7	36.8	168	2 T34804	hypothetical prote
22	7	36.8	219	1 XMECAD	deda protein - Esc
23	7	36.8	219	2 A98029	hypothetical prote
24	7	36.8	229	2 B35173	2,2-dialkylglycine
25	7	36.8	231	2 F75459	hypothetical prote
26	7	36.8	248	2 T35233	hypothetical prote
27	7	36.8	255	2 T36243	probable precorrin
28	7	36.8	256	2 T03371	glycine-rich prote
29	7	36.8	276	2 T43203	probable alpha/alp

30 7 36.8 292 2 JE0233 troponin-I - scall  
31 7 36.8 307 2 T45274 transcription fact  
32 7 36.8 315 2 G87464 nitrogen regulatio  
33 7 36.8 316 2 B70620 hypothetical prote  
34 7 36.8 323 2 E86935 conserved hypothet  
35 7 36.8 324 2 A87084 conserved hypothet  
36 7 36.8 326 2 B87557 hypothetical prote  
37 7 36.8 326 2 S27534 hypothetical prote  
38 7 36.8 338 2 D85851 probable oxidoredu  
39 7 36.8 339 2 F83026 conserved hypothet  
40 7 36.8 344 2 H70710 probable hemz prot  
41 7 36.8 353 2 T35671 probable transmemb  
42 7 36.8 354 2 B57333 G protein-coupled  
43 7 36.8 360 2 F70819 probable ionicttran  
44 7 36.8 383 2 T35773 succinyl-CoA synth  
45 7 36.8 393 2 D70858 probable iron-sulf

## ALIGNMENTS

## RESULT 1

JC5700

ErbB kinase activator alpha, brain and thymus - human

C:Species: Homo sapiens (man)

C:Date: 25-Nov-1997 #sequence\_revision 25-Nov-1997 #text\_change 08-Sep-2002

C:Accession: JC5700

R:Higashiyama, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.;

J. Biochem. 122, 675-680, 1997

A:Title: A novel brain-derived member of the epidermal growth factor family that i

A:Reference number: JC5700; MUID:98006324; PMID:9348101

A:Accession: JC5700

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-850 <HIG>

A:Cross-references: DBJ:AB005060; MID:g2626738; PIDN:BAA23417.1; PID:g2626739

A:Experimental source: SK-NSH cell

C:Comment: This protein is a member of the epidermal growth factor family. It is f

ating the differentiation of MDA-MB-453 cells.

C:Superfamily: human ErbB kinase activator alpha, brain and thymus; EGF homology; f

C:Keywords: glycoprotein

F:258-381/Domain: Ig-like #status predicted <IGL>

F:345-381/Domain: EGF-like #status predicted <EGF>

F:346-381/Domain: EGF-like #status predicted <EGF2>

F:147,278,451/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 47.4%; Score 9; DB 2; Length 850;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EARRARRAA 10

DB 554 EARRARRAA 562

## RESULT 2

JC5702

ErbB kinase activator alpha2a, brain and thymus - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 25-Nov-1997 #sequence\_revision 25-Nov-1997 #text\_change 08-Sep-2002

C:Accession: JC5702; PC4417

R:Higashiyama, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.;

J. Biochem. 122, 675-680, 1997

A:Title: A novel brain-derived member of the epidermal growth factor family that in

A:Reference number: JC5700; MUID:98006324; PMID:9348101

A:Accession: JC5702

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-860 <HIG>

A:Cross-references: DBJ:D89996; MID:g2605631; PIDN:BAA23345.1; PID:g2605632

A:Experimental source: PC-12 cell

A:Accession: PC4417

A:Status: nucleic acid sequence not shown

DATE: 24 AUG 2002 FREQUENCY: 24 AUG 2002 15Z  
 2002-SEP-30 0001 15Z LCHANNP 30-SEP-2002

C:Accession: E95279  
 R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
 A:Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*  
 A:Reference number: A95262; MUID:21396509; PMID:11481432  
 A:Accession: E95279  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-334 <KUR>  
 A:Cross-references: GB:AE006469; PIDN:AAK64799.1; PID:g14523209; GSPDB:GN00165  
 A:Experimental source: strain 1021, megaplasmid pSymA  
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: SWA0265  
 A:Genome: plasmid  
 C:Superfamily: malate dehydrogenase ylbC  
 Query Match 42.1%; Score 8; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 ARARRAAA 10  
 DB 306 ARARRAAA 313  
 RESULT 8  
 KISMG  
 galactokinase (EC 2.7.1.6) - Streptomyces sp.  
 C:Species: Streptomyces sp.  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 28-May-1999  
 C:Accession: C28669  
 R:Adams, C.W.; Fornwald, J.A.; Schmidt, F.J.; Rosenberg, M.; Brawner, M.E.  
 J. Bacteriol. 170, 203-212, 1988  
 A:Title: Gene organization and structure of the *Streptomyces lividans* gal operon.  
 A:Reference number: A28669; MUID:88086869; PMID:3335481  
 A:Accession: C28669  
 A:Molecule type: DNA  
 A:Residues: 1-395 <ADA>  
 A:Cross-references: GB:M18953; NID:g153259; PIDN:AAA26748.1; PID:g153262  
 A:Note: the source is designated as *Streptomyces lividans*  
 C:Genetics:  
 A:Gene: galK  
 C:Superfamily: galactokinase  
 C:Keywords: ATP; galactose metabolism; phosphotransferase  
 Query Match 42.1%; Score 8; DB 1; Length 395;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 AARARRA 16  
 DB 228 AARARRA 235  
 RESULT 9  
 T34741  
 hypothetical protein SC1E6.20c SC1E6.20c - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T34741  
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, November 1998  
 A:Reference number: Z21555

A:Accession: T34741  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-501 <SEE>  
 A:Cross-references: EMBL:AL033505; PIDN:CAA22048.1; GSPDB:GN00070; SCOEDB:SC1E6.20c  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC1E6.20c  
 Query Match 42.1%; Score 8; DB 2; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RARRAAA 11  
 DB 71 RARRAAA 78  
 RESULT 10  
 E87343  
 ABC transporter, ATP-binding protein CydC CC0760 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
 C:Accession: E87343  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: E87343  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-532 <STO>  
 A:Cross-references: GB:AE005673; NID:g13421995; PIDN:AAK22745.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC0760  
 C:Superfamily: Mycobacterium tuberculosis probable ABC transporter cydC; ATP-binding  
 Query Match 42.1%; Score 8; DB 2; Length 532;  
 Best Local Similarity 100.0%; Pred. No. 9.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 ARRAARA 12  
 DB 243 ARRAARA 250  
 RESULT 11  
 S06188  
 genome polyprotein 1 - grapevine chrome mosaic virus  
 N:Alternate names: RNA1 polyprotein  
 N:Contains: cysteine proteinase (EC 3.4.22.-); glycoprotein VPg; nucleotide-binding m  
 C:Species: grapevine chrome mosaic virus  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Sep-1999  
 C:Accession: S06188  
 R:Le Gall, O.; Candresse, T.; Brault, V.; Dunez, J.  
 Nucleic Acids Res. 17, 7795-7807, 1989  
 A:Title: Nucleotide sequence of Hungarian grapevine chrome mosaic nepovirus RNA1.  
 A:Reference number: S06188; MUID:90016864; PMID:2798128  
 A:Accession: S06188  
 A:Molecule type: mRNA  
 A:Residues: 1-2252 <LEG>  
 A:Cross-references: EMBL:X15346; NID:g59345; PIDN:CAA33405.1; PID:g59346  
 A:Note: 1353-Arg and 1870-Arg were also found  
 C:Note: the sequence translated from one cDNA differs from that shown in having 1571-  
 C:Genetics:  
 A:Map position: segment 1  
 C:Superfamily: cowpea aphid-borne mosaic virus genome polyprotein B  
 C:Keywords: cysteine proteinase; genome-linked protein; hydrolase; membrane protein;  
 Query Match 42.1%; Score 8; DB 2; Length 2252;  
 Best Local Similarity 100.0%; Pred. No. 31;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARRAAAR 11  
|||||

Db 249 RARRAAAR 256

# RESULT 12

G84072  
hypothetical protein BH3383 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G84072  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G84072  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-57 <STO>  
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA07102.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3383

Query Match 36.8%; Score 7; DB 2; Length 57;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AARRAAR 14

|||||

Db 27 AARRAAR 33

# RESULT 13

AI3052  
hypothetical protein Atu4038 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AI3052  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AI3052  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-71 <KOR>  
A:Cross-references: GB:AE008689; PIDN:AAL44839.1; PID:g17742483; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu4038  
A:Map position: linear chromosome

Query Match 36.8%; Score 7; DB 2; Length 71;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RARRAAAR 17

|||||

Db 39 RARRAAAR 45

# RESULT 14

A87663  
hypothetical protein CC3339 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: A87663

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87663

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-91 <STO>

A:Cross-references: GB:AE005673; NID:g13425039; PIDN:AAK25301.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3339

Query Match 36.8%; Score 7; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AARRAAR 15

|||||

Db 81 AARRAAR 87

# RESULT 15

T52525  
tyrosinase co-factor MelC2 [imported] - Streptomyces galbus  
C:Species: Streptomyces galbus  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 24-Oct-2000  
C:Accession: T52525  
R;Wehmeier, U.F.; Brass, N.; Roessler, C.; Piepersberg, W.  
submitted to the EMBL Data Library, February 1996  
A:Description: Cloning and characterization of the mel-operon from Streptomyces galbu  
A:Reference number: Z26098  
A:Accession: T52525  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-126 <WEH>  
A:Cross-references: EMBL:X95705; PIDN:CAA65004.1  
A:Experimental source: DSM40480  
C:Genetics:  
A:Gene: melC1

Query Match 36.8%; Score 7; DB 2; Length 126;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RARRAAAR 13

|||||

Db 105 RARRAAAR 111

Search completed: August 9, 2003, 16:34:11

Job time : 19.4571 secs